

The validity of general practitioners' self assessment of knowledge: cross sectional study

Jocelyn M Tracey, Bruce Arroll, David E Richmond, Philip M Barham

Goodfellow Unit,
Faculty of Medicine
and Health Science,
University of
Auckland, Private
Bag 92019,
Auckland,
New Zealand

Jocelyn Tracey,
assistant director
Philip Barham,
director

Department of
General Practice,
Faculty of Medicine
and Health Science,
University of
Auckland

Bruce Arroll,
associate professor

Faculty of Medicine
and Health Science,
University of
Auckland

David Richmond,
dean of
undergraduate studies

Correspondence to:
Dr Tracey
j.tracey@auckland.
ac.nz

BMJ 1997;315:1426-8

Abstract

Objective: To determine whether general practitioners can make accurate self assessments of their knowledge in specific areas.

Design: 67 general practitioners completed a self assessment of their level of knowledge over a variety of topics using a nine point semantic differential scale. An objective assessment of their knowledge was then made by administering true-false tests on two of the topics: thyroid disorders and non-insulin dependent diabetes. The study was repeated with another group of 60 general practitioners, using sexually transmitted diseases as the topic.

Setting: General practices in New Zealand.

Subjects: Random sample of 67 general practitioners in Auckland.

Main outcome measure: Test scores for self assessment and for actual knowledge.

Results: Correlations between self assessments and test scores were poor for all three topics studied ($r=0.19$ for thyroid disorders, 0.21 for non-insulin dependent diabetes, 0.19 for sexually transmitted diseases).

Conclusions: As general practitioners cannot accurately assess their own level of knowledge on a given topic, professional development programmes that rely on the doctors' self perceptions to assess their needs are likely to be seriously flawed.

Introduction

Most continuing medical education programmes follow the theories of self directed learning expounded by Rogers¹ and Knowles.²⁻³ There is some evidence that learning programmes are more effective if they are based on the participants' learning needs.⁴⁻⁵ Accordingly, most organisers of continuing medical education assess the learning needs of participants before designing the programmes.⁶ Many recertification programmes are also based on the principles of self directed learning, expecting participants to select for themselves educational activities appropriate to their needs.⁷

We found only one published study that assessed the accuracy of general practitioners' self assessments.⁸ It showed a poor correlation between generalised self assessment of proficiency of skills and the results from performance based tests. We set out to determine whether general practitioners can make accurate self assessments of their knowledge in specific areas.

Methods

A sample of 100 general practitioners was selected from a database of current general practitioners in Auckland by using a random numbers chart. The general practitioners were mailed a coded questionnaire asking them to rate on a nine point semantic differential scale their level of knowledge of 20 problems which present in a typical general practice. The extremities of the scale were characterised by large gaps in knowledge at one pole and comprehensive knowledge at the other. Participants were asked to evaluate the level of their knowledge without resource to texts or journals.

After the questionnaire results were analysed, thyroid disorders and non-insulin dependent diabetes mellitus were chosen for further testing because the self assessment scores for these topics formed a normal bell curve with a wide spread.

A written true-false test of 50 items on thyroid disorders was developed on the basis of a review of the literature and refined with the assistance of two endocrinologists and seven active general practitioners (box). Only questions rated as clinically relevant to general practice by the panel general practitioners were included. Distribution of question items reflected the frequency of the range of thyroid problems in general practice.

Tests were mailed to participants, who completed them in their own time without studying, recourse to literature, or the help of colleagues. The time delay between self assessment and test was 3 months. There were no continuing medical education courses or readily available journal articles on thyroid disorders for general practitioners during this time.

Examples of true-false test questions

A raised TSH is the least sensitive marker of decreased thyroid reserve in Hashimoto's thyroiditis (False)

T4 treatment gives useful shrinkage in the treatment of euthyroid diffuse goitre (True)

An elderly patient with hyperthyroidism may only have symptoms and signs of general debility (True)

Most patients with toxic nodular goitre become hypothyroid after radioiodine treatment (False)

Antimicrosomal antibodies are usually present in primary hypothyroidism (True)

Pearson's correlation coefficient was used to assess the relation between the participants' self assessments of knowledge and their scores on the written test. The internal consistency of the test was analysed by using Cronbach's coefficient α . Test items that correlated poorly with the rest of the test were removed, and the correlation between self assessments and validated test scores was recalculated.

This procedure was repeated with non-insulin dependent diabetes as the test topic, and was later repeated a third time to exclude three of the uncontrolled variables: time delay, examination conditions, and lack of complete anonymity. The topic chosen was sexually transmitted diseases. The test was run at the beginning of a continuing medical education course to avoid a time delay between the self assessment and the test; the examination conditions were uniform for the group; and there was no threat of identification of test results.

Results

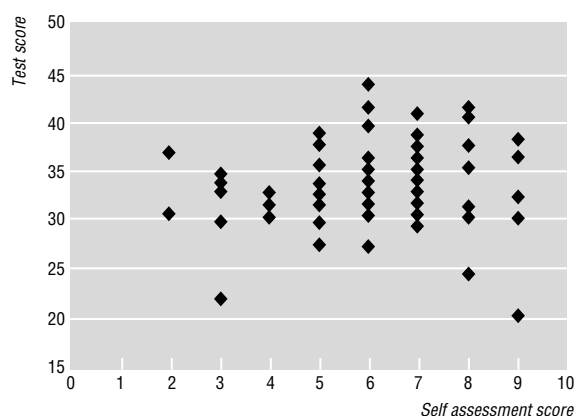
Self assessment questionnaire

Of the 100 participants, 73 returned their self assessment questionnaires. The only significant difference between responders and non-responders was that non-members of the Royal New Zealand College of General Practitioners were less likely to participate in the study ($P=0.0009$, t test). Sex, size of practice, and year of registration had no effect.

The ratings by the general practitioners on the different topics showed a variety of patterns of response. Both thyroid disorders and non-insulin dependent diabetes had a wide spread with a normal bell shaped curve. The mean self assessment score for thyroid disorders was 6.06 (SD 1.7; range 2-9). Mean self assessment score for non-insulin dependent diabetes was 6.84 (1.7; 1-8).

Other topics, especially conditions that are common in general practice, such as paediatric asthma and contraception, had frequency distributions skewed to the right, indicating a greater confidence in level of knowledge.

Self assessment scores over all 20 topics had a mean of 6.46 (1.10; mode = 4). Most general practitioners used a wide range of points on the scale, indicating that they perceived strengths in some areas and weaknesses in others.



Correlation between general practitioners' self assessment of and test scores for knowledge of thyroid disorders

Correlations between self assessment and test scores

Topic	Initial correlation	Correlation after exclusion of weak test items
Thyroid disorders	0.19	0.19
Non-insulin dependent diabetes	0.12	0.21
Sexually transmitted diseases	0.20	0.19

Thyroid test

Two thirds of the total study population (67 general practitioners; 92% of those who completed the self assessment questionnaire) completed the thyroid test. Of a possible score of 50, the mean test score was 34.2 (4.61; 21-44; mode = 33). The correlation between the self assessment and thyroid test scores was 0.195 (figure).

Internal consistency analysis gave a Cronbach coefficient α of 0.63 (0.67 after removal of five items which correlated negatively with total score). With this statistically more reliable test, the correlation between the self assessment scores and test scores was no better (0.194 (95% confidence interval -0.049 to 0.415), $P=0.11$).

As numbers were not sufficient to stratify according to sex, practice size, age, and membership of royal college, separate correlations between self assessment and test scores were calculated for each subgroup and tested using Fisher's z transformation. No significant differences were found, suggesting that these variables had little effect on the strength of the relation.

Diabetes

A total of 56 general practitioners completed the diabetes test. The mean score was 33.9 out of a possible score of 45; the correlation between the diabetes self assessment scores and test results was 0.117. Reliability analysis gave a Cronbach coefficient of 0.31, which was improved to 0.60 by removal of poor test items. The correlation between self assessment scores and new test scores was 0.206 ($P=0.12$).

Sexually transmitted diseases

Of the 75 attendees at the continuing medical education course, 60 completed the self assessment score and 70 completed the true-false test. The mean self assessment score was 5.11 (1.80; 2-8); the mean score on the test was 23.95 out of a possible 35. The correlation between self assessment scores and test results was 0.20. After removal of test items to improve the validity of the test the correlation was 0.19 ($P=0.15$).

Correlations between the self assessment scores and test scores were similar for all three tests (table).

Discussion

This research shows that general practitioners' insight into their educational needs is poor. Correlations were uniformly low between self assessment scores and test scores over three different topics representative of the variety of conditions which general practitioners are likely to see. Some participants who thought they were knowledgeable in an area were shown not to be; others who thought their knowledge was poor scored well on testing.

Key messages

- Doctors' perception of knowledge in areas of common practice is no indication of actual knowledge
- Continuing medical education and other professional development activities that rely on the doctors' self perception to assess their needs are likely to be seriously flawed
- To make professional development activities more efficient and effective a more objective assessment of needs is necessary

Methodological concerns

As semantic differential scales are commonly used to evaluate medical education activities, general practitioners would be expected to use our self assessment scale appropriately. It was unhelpful to measure the internal consistency of the scale because the 20 topics had no relation to each other. However, the results obtained from the scale suggest that it was used appropriately. The frequency distributions obtained for most topics showed normal bell shaped curves and the more common topics were skewed to the right, as expected. Each general practitioner's responses varied from topic to topic, with most using a wide range of responses.

The methods used to develop the true-false tests provide some assurance of content and face validity. The Cronbach α correlation coefficients of 0.67 for thyroid, 0.60 for diabetes and 0.47 were acceptable,⁹ especially considering that a group of general practitioners with a wide variety of experience would not be expected to have a homogeneous knowledge of these areas. They also compare favourably with the reliability coefficient of 0.43 for the written test used in a similar study.⁸

The group that completed the first two tests was chosen at random from all general practitioners in Auckland. Response rates were reasonable. Non-members of the Royal New Zealand College of General Practitioners, a group that might have been expected to be less self aware because of their lack of involvement in any ongoing accreditation, were over-represented in the non-responder group. Although the general practitioners who were tested on sexually transmitted diseases were not chosen at random, their inclusion allowed potential sources of error present for the first two topics to be controlled.

Perceived and actual knowledge

Jansen et al, comparing the accuracy of general practitioners' self assessments of technical clinical skills with evaluation on a performance based test and a written test, found correlations of 0.30 and 0.29, respectively (similar to the results of our study).⁸ The reliability coefficients of both their tests were quite low at 0.43. When they recalculated their correlations after correction for the attenuation caused by the unreliability of the tests, the correlations improved to 0.47 and 0.49 respectively. They used statistical methods to artificially increase the reliability coefficient of the tests to 0.80, rather than deleting those test items with low Pearson correlation coefficients. Using the more rigorous method to improve the reliability of the three tests in our study failed to improve the correlations.

Jansen et al summed their assessments over eight quite different topics. The general practitioners' generalised self assessments of competence were then corre-

lated with their knowledge and skills over the same broad area. In our study, the general practitioners' self assessment of one specific, defined area was correlated with knowledge in that defined area. General practitioners' self assessment of general competence may be slightly more accurate than their self assessment of knowledge in a specific area.

The lack of correlation between perceived and actual knowledge is not congruent with the principles of self directed learning on which most continuing medical education programmes are based.^{1-3 10} It suggests that general practitioners are not able accurately to assess their specific learning needs, and hence that self directed learning activities may be misdirected.

Two reasons may be advanced to explain the difficulty general practitioners have in assessing their knowledge. The isolationist nature of their work means that there is little opportunity for regular discussion with peers on clinical matters, and the breadth of general practice in a time of rapidly changing knowledge may make it difficult for general practitioners to be aware of where their knowledge falls short.

Implications

This research has important implications for continuing medical education and for the style of recertification programmes which have been developed in Australasia. Educational programmes based on self directed learning and continuous quality improvement are good in theory and are a valuable method of learning, but if the initial self assessment of learning needs is faulty, then the system is fundamentally flawed.

An ideal recertification programme should "ensure that the physician has the potential to respond appropriately to the wide range of problems that do not occur often enough in practice to be evaluated in his or her patients' outcomes."¹¹ As many gaps in clinical knowledge will not be uncovered by performance measures, a diagnostic written test of knowledge and clinical judgment could be an integral part of recertification programmes.¹¹

The results of our study strongly support the need for better self assessment tools to help identify learning needs for continuing medical education and recertification programmes.

Funding: New Zealand Lottery Grants Board.

Conflict of interest: None.

- 1 Rogers CR. *Freedom to learn*. Columbus, OH: Merrill, 1969.
- 2 Knowles MS. *The adult learner: a neglected species*. 2nd ed. Houston: Gulf, 1978.
- 3 Knowles MS. *Andragogy in action: applying modern principles of adult learning*. San Francisco: Jossey-Bass, 1984.
- 4 Stein LS. The effectiveness of continuing medical education: eight research reports. *J Med Educ* 1981;56:103-10.
- 5 Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. *JAMA* 1995;274:700-5.
- 6 Ward J. Needs assessment in continuing medical education. *Med J Aust* 1988;148:77-80.
- 7 Gabb R. Recertification: an Australasian perspective. In: Newble D, Jolly B, Wakeford R, eds. *The certification and recertification of doctors*. Cambridge: Cambridge University Press, 1994:47-66.
- 8 Jansen JJM, Tan LHC, van der Vleuten CPM, van Luijk SJ, Rethans JJ, Grol RPTM. Assessment of competence in technical clinical skills of general practitioners. *Med Educ* 1995;29:247-53.
- 9 SAS technical tips: measuring the internal consistency of a test. *SAS Communications* 1994;20:1-5.
- 10 Woolliscroft JO, TenHaken J, Smith J, Calhoun JG. Medical students' clinical self-assessments: comparisons with external measures of performance and the student's self assessments of overall performance and effort. *Acad Med* 1993;68:285-94.
- 11 Norcini JJ. Recertification in the medical specialties. *Acad Med* 1994;69:S90-3.

(Accepted 11 August 1997)